Medidas de dispersão

Professor Fiore

As medidas de dispersão são importantes, pois indica o quão ‘espalhados’ os valores estão. As principais medidas são a amplitude total, a variância, o desvio-padrão e o coeficiente de variação.

Amplitude total

A amplitude total (A_T) é a diferença entre o maior e o menor valor da série.

10 13 15 23 24

$A_T = 24 – 10 = 14$

Variância

A variância é uma medida que considera todos os valores e para encontra-la basta seguir os passos abaixo.

1. Identifique se os valores representam toda população ou apenas uma amostra. Quando não houver a informação, considere os valores como amostra.
2. Organize os dados em uma tabela (horizontal ou vertical).

Dados	x_i	10	13	15	23	24

3. Calcule a média dos valores $\bar{x} = \frac{85}{5} = 17$
4. Registre em uma nova linha (ou coluna) os valores dos desvios, a diferença entre cada valor e a média $(x_i - \bar{x})$.

Dados	x_i	10	13	15	23	24
Desvios	$(x_i - \bar{x})$	-7	-4	-2	6	7

5. Registre na próxima linha (ou coluna) o valor dos desvios elevado ao quadrado, $(x_i - \bar{x})^2$.

Dados	x_i	10	13	15	23	24
Desvios	$(x_i - \bar{x})$	-7	-4	-2	6	7
Desvios ao quadrado	$(x_i - \bar{x})^2$	49	16	4	36	49

6. A variância amostral é a soma dos valores dos desvios ao quadrado, soma dos valores da última linha (ou coluna), dividido pelo número de valores menos um.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{154}{5-1} = 38{,}5$$

Se os dados representam uma população, o denominador será o valor de n, não subtraímos um.

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

Desvio padrão

O desvio padrão é a raiz quadrada da variância, amostral para amostra e populacional para população, conforme visto anteriormente. Para o desvio padrão amostral o símbolo s é usado e para populacional o símbolo σ.

$$Desvio \;\; Padrão = \sqrt{Variância} \qquad s = \sqrt{s^2} \qquad \sigma = \sqrt{\sigma^2}$$

Considerando os dados do exemplo anterior, temos o desvio padrão amostral $s = \sqrt{38{,}5} = 6{,}20$.

Coeficiente de variação

O coeficiente de variação (C_v) é a razão entre o desvio padrão e a média, geralmente escrita como porcentagem.

$$C_v = \frac{Desvio \;\; Padrão}{Média}$$

O símbolo segue ser de acordo com os dados, para população escrevemos $C_v = \frac{\sigma}{\mu}$ e para amostra $C_v = \frac{s}{\bar{x}}$.

No exemplo temos: $C_v = \frac{s}{\bar{x}} = 0{,}36 = 36\%$

Resumo:

	População	**Amostra**
Média	$\mu = \frac{\sum x_i}{N}$	$\bar{x} = \frac{\sum x_i}{n}$
Variância	$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$
Desvio padrão	$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$	$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

Apesar da diferença entre o símbolo para população e amostra, o processo de cálculo só difere para a variância e para o desvio padrão, raiz da variância.

Na maioria das vezes lidamos com amostras, sendo assim o mais comum é calcular as medidas amostrais.